Central European Journal of Economic Modelling and Econometrics

Volatility Persistence and Predictability of Squared Returns in GARCH(1,1) Models

Umberto Triacca*

Submitted: 1.07.2009, Accepted: 6.01.2010

Abstract

Volatility persistence is a stylized statistical property of financial time-series data such as exchange rates and stock returns. The purpose of this letter is to investigate the relationship between volatility persistence and predictability of squared returns.

Keywords: GARCH Models, returns, time series, volatility persistence

JEL Classification: C13, C22, G11, G12.

*University of L'Aquila, Rome, email: umberto.triacca@gmail.com

1 Introduction

The one-period return on a stock with price P_t at time t is defined as

$$y_t = \log(P_t) - \log(P_{t-1}).$$

Let $\{\mathcal{F}_t\}$ be a filtration (an increasing sequence of sigma algebras) modeling the information set available at time t. We assume

$$y_t = \sigma_t z_t \tag{1}$$

where $z_t \sim i.i.d.(0,1)$ and adapted to $\{\mathcal{F}_t\}$ and σ_t is a stochastic process adapted to $\{\mathcal{F}_{t-1}\}$. The process $\{x_t\}$ is said to be adapted to the filtration $\{\mathcal{F}_t\}$ if for each $t \geq t_0$, x_t is $\mathcal{F}_t$-measurable.
We have $E(y_t|\mathcal{F}_{t-1}) = 0$ and $E(y_t^2|\mathcal{F}_{t-1}) = \sigma_t^2$. The process $\{y_t\}$ has conditional mean zero and it is conditionally heteroskedastic with conditional variance σ_t^2. Thus σ_t represents the volatility of the price change between times $t-1$ and t.
Volatility persistence is a stylized statistical property of financial time-series data such as exchange rates and stock returns. The purpose of this note is to investigate the relationship between volatility persistence and predictability of squared returns, y_t^2.

2 The result

In order to explicitly take into account volatility persistence in the returns series, we assume that y_t follows a GARCH(1,1) model. It provides a measure of volatility expressed as follows:

$$\sigma_t^2 = \omega + \alpha_1 y_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{2}$$

where ω, α_1, and β_1 are parameters such that $\omega > 0$, α_1, $\beta_1 \geq 0$.
We shall make the following two assumptions: (A.1) $\alpha_1 + \beta_1 < 1$ (A.2) $(\alpha_1 + \beta_1)^2 + \alpha_1^2(\kappa_z - 1) < 1$, where κ_z is the kurtosis of z_t.
The coefficients α_1 and β_1 reflect the dependence of the current volatility upon its past levels and the sum $\alpha_1 + \beta_1$ indicates the degree of volatility persistence. To see this we rewrite equation (2) as

$$\sigma_t^2 = \omega + (\alpha_1 + \beta_1)\sigma_{t-1}^2 + \alpha_1 \nu_{t-1}$$

where $\nu_{t-1} = y_{t-1}^2 - \sigma_{t-1}^2$. It follows that

$$\sigma_t^2 = \frac{\omega}{1 - \alpha_1 - \beta_1} + \alpha_1 \left[\nu_{t-1} + (\alpha_1 + \beta_1)\nu_{t-2} + (\alpha_1 + \beta_1)^2 \nu_{t-3} + ...\right] \tag{3}$$

Equation (3) shows that $\alpha_1 + \beta_1$ determines how long a random shock to volatility persists. Thus the sum $\phi = \alpha_1 + \beta_1$ is often referred to as the persistence parameter.

Now, we consider a measure of predictability of the squared returns, y_t^2, relative to h-steps forecast defined by

$$R^2(h) = 1 - \frac{\text{var}(e_t(h))}{\text{var}(y_t^2)}$$

where $e_t(h) = y_{t+h}^2 - E(y_{t+h}^2|\mathcal{F}_t)$. This predictability index has been utilized also by Hong and Billings (1999), Otranto and Triacca (2007) and Pena and Sanchez (2007). We observe that in the ARCH(1) case (i.e. $\beta_1 = 0$) we have

$$R^2(h) = \alpha_1^{2h}, \qquad h = 1, ...$$

Thus it is trivial to conclude that:

1. $\alpha_1 = \sqrt{\frac{R^2(h+1)}{R^2(h)}}$

2. $\lim_{h\to\infty} \sqrt[2h]{R^2(h)} = \alpha_1$

In this note we will show that this results hold also for a GARCH(1,1) model.
We first show that

$$R^2(h) \quad = \quad \frac{\alpha_1^2(\alpha_1+\beta_1)^{-2}(\alpha_1+\beta_1)^{2h}}{1-2\alpha_1\beta_1-\beta_1^2}$$

In order to do this, we rewrite the equation for σ_t^2 in (2) with $\nu_t = y_t^2 - \sigma_t^2$, obtaining the following well-known ARMA(1,1) representation for that y_t^2:

$$y_t^2 = \omega + \phi y_{t-1}^2 + \nu_t - \beta_1\nu_{t-1} \tag{4}$$

The equation (4) can be written in the more compact form

$$\phi(B)y_t^2 = \omega + \beta_1(B)\nu_t \tag{5}$$

where B is the backward shift operator, $\phi(B) = 1 - \phi B$ and $\beta_1(B) = 1 - \beta_1 B$. Under assumption (A.1), the ARMA representation (5) is causal and invertible (although $\sigma_\nu^2 = E(\nu_t^2)$ is not necessarily finite). The assumptions (A.1) and (A.2) ensure that $\sigma_\nu^2 < \infty$.
By section 3.1 of Brockwell and Davis (1991), causality implies that there exists a sequence of constants $\{\psi_i\}$ such that

$$\sum_{j=0}^{\infty} |\psi_j| < \infty$$

and

$$y_t^2 = \sum_{j=0}^{\infty} \psi_j \nu_{t-j} + \mu \quad t = 0, \pm 1, ...$$

The ψ_j's are obtained from the relation

$$\psi(z)\phi(z) = \beta_1(z)$$

with $\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j \quad |z| < 1$.
In particular, we have $\psi_0 = 1$ and $\psi_j = \alpha_1(\alpha_1 + \beta_1)^{j-1}$ for $j \geq 1$. Thus

$$\begin{aligned}
\sum_{j=0}^{\infty} \psi_j^2 &= 1 + \alpha_1^2 + \alpha_1^2(\alpha_1+\beta_1)^2 + \alpha_1^2(\alpha_1+\beta_1)^4 ... \\
&= 1 + [1 + (\alpha_1+\beta_1)^2 + (\alpha_1+\beta_1)^4 + ...]\alpha_1^2 \\
&= 1 + \left[\frac{1}{1-(\alpha_1+\beta_1)^2}\right]\alpha_1^2 \\
&= \frac{1-2\alpha_1\beta_1-\beta_1^2}{1-(\alpha_1+\beta_1)^2}
\end{aligned}$$

and

$$\begin{aligned}
\sum_{j=0}^{h-1} \psi_j^2 &= \sum_{j=0}^{\infty} \psi_j^2 - \sum_{j=h}^{\infty} \psi_j^2 \\
&= \frac{1-2\alpha_1\beta_1-\beta_1^2}{1-(\alpha_1+\beta_1)^2} - [\alpha_1^2(\alpha_1+\beta_1)^{2(h-1)} + \alpha_1^2(\alpha_1+\beta_1)^{2h} ...] \\
&= \frac{1-2\alpha_1\beta_1-\beta_1^2}{1-(\alpha_1+\beta_1)^2} - [1 + \alpha_1^2(\alpha_1+\beta_1)^2 + ...]\alpha_1^2(\alpha_1+\beta_1)^{2(h-1)} \\
&= \frac{1-2\alpha_1\beta_1-\beta_1^2}{1-(\alpha_1+\beta_1)^2} - \frac{\alpha_1^2(\alpha_1+\beta_1)^{2(h-1)}}{1-(\alpha_1+\beta_1)^2}
\end{aligned}$$

and hence we have

$$\begin{aligned}
\mathrm{var}(y_t^2) &= (1 + \psi_1^2 + ...)\sigma_\nu^2 \\
&= \left[\frac{1-2\alpha_1\beta_1-\beta_1^2}{1-(\alpha_1+\beta_1)^2}\right]\sigma_\nu^2
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{var}(e_t(h)) &= (1 + \psi_1^2 + ... + \psi_{h-1}^2)\sigma_\nu^2 \\
&= \left[\frac{1-2\alpha_1\beta_1-\beta_1^2}{1-(\alpha_1+\beta_1)^2} - \frac{\alpha_1^2(\alpha_1+\beta_1)^{2(h-1)}}{1-(\alpha_1+\beta_1)^2}\right]\sigma_\nu^2
\end{aligned}$$

It follows that

$$\begin{aligned} R^2(h) &= 1 - \frac{1-2\alpha_1\beta_1-\beta_1^2-\alpha_1^2(\alpha_1+\beta_1)^{2(h-1)}}{1-2\alpha_1\beta_1-\beta_1^2} \\ &= \frac{\alpha_1^2(\alpha_1+\beta_1)^{2(h-1)}}{1-2\alpha_1\beta_1-\beta_1^2} \\ &= \frac{\alpha_1^2(\alpha_1+\beta_1)^{-2}(\alpha_1+\beta_1)^{2h}}{1-2\alpha_1\beta_1-\beta_1^2} \end{aligned}$$

Now, we can show that the persistence parameter $\phi = \alpha_1 + \beta_1$ can be expressed in terms of the predictability's measure of squared returns. We have

$$\begin{aligned} R^2(h+1) &= \frac{\alpha_1^2(\alpha_1+\beta_1)^{2(h+1-1)}}{1-2\alpha_1\beta_1-\beta_1^2} \\ &= \frac{\alpha_1^2(\alpha_1+\beta_1)^{2(h-1)}(\alpha_1+\beta_1)^2}{1-2\alpha_1\beta_1-\beta_1^2} \\ &= R^2(h)(\alpha_1+\beta_1)^2 \end{aligned}$$

Thus

$$\alpha_1 + \beta_1 = \sqrt{\frac{R^2(h+1)}{R^2(h)}}$$

We conclude this section obtaining the persistence parameter $\phi = \alpha_1 + \beta_1$ as limit of the sequence $\left\{ \sqrt[2h]{R^2(h)} \right\}$.
We have

$$\begin{aligned} \lim_{h\to\infty} \sqrt[2h]{R^2(h)} &= \lim_{h\to\infty} \sqrt[2h]{\frac{\alpha_1^2(\alpha_1+\beta_1)^{-2}(\alpha_1+\beta_1)^{2h}}{1-2\alpha_1\beta_1-\beta_1^2}} \\ &= (\alpha_1+\beta_1) \lim_{h\to\infty} \sqrt[2h]{\frac{\alpha_1^2(\alpha_1+\beta_1)^{-2}}{1-2\alpha_1\beta_1-\beta_1^2}} \\ &= \alpha_1+\beta_1 \\ &= \phi \end{aligned}$$

3 A simulation study

In this paper we have investigated the relationship between the GARCH(1,1) persistence parameter ϕ and the R^2 of h-step forecasts of squared returns. In particular we have shown that the persistence parameter ϕ can be obtained as limit of the sequence $\left\{ \sqrt[2h]{R^2(h)} \right\}$. As an illustration of how this analytic relationship can be used in the practice, we note that if the maximum likelihood estimation (MLE) of ϕ, $\hat{\phi} = \hat{\alpha}_1 + \hat{\beta}_1$, is downward biased and if

$$\frac{\hat{\alpha}_1^2(\hat{\alpha}_1 + \hat{\beta}_1)^{-2}}{1 - 2\hat{\alpha}_1\hat{\beta}_1 - \hat{\beta}_1^2} > 1$$

then there exists a $\delta \in \mathbb{N}$ such that the estimator

$$\sqrt[2h]{\hat{R}^2(h)} = \left(\hat{\alpha}_1 + \hat{\beta}_1\right) \sqrt[2h]{\frac{\hat{\alpha}_1^2(\hat{\alpha}_1 + \hat{\beta}_1)^{-2}}{1 - 2\hat{\alpha}_1\hat{\beta}_1 - \hat{\beta}_1^2}}$$

for $h \geq \delta$ produces parameter estimates which compare favorably with that of the MLE.

This fact is relevant since it is well known that the MLE of ϕ is often severely downward biased in small samples; see Bollerslev, Engle, Nelson (1994) and Hwang and Valls Pereira (2006).

In order to show how the estimator $\sqrt[2h]{\hat{R}^2(h)}$ works a small Monte Carlo experiment is conducted. The simulation results are obtained with 1000 replications for the following GARCH(1,1) model:

$$y_t = \sigma_t z_t$$

$$\sigma_t^2 = \omega + \alpha_1 y_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

with $\omega = 0.01$, $\alpha_1 = 0.2$, $\beta_1 = 0.6$ (DGP I) and with $\omega = 0.01$, $\alpha_1 = 0.1$, $\beta_1 = 0.6$ (DGP II). These values are utilized also in the simulation experiment presented in Hwang and Valls Pereira (2006). When the DGP I is used and the sample size is 100, in the 78.9% of cases the estimator $\sqrt[14]{\hat{R}^2(7)}$ (we have posed $h = 7$) performs better than MLE $\hat{\phi}$. When the DGP II is used and the sample size is 100, this percentage rises to the 88.8%.

The results from our Monte Carlo study suggest that when

$$\frac{\hat{\alpha}_1^2(\hat{\alpha}_1 + \hat{\beta}_1)^{-2}}{1 - 2\hat{\alpha}_1\hat{\beta}_1 - \hat{\beta}_1^2} > 1$$

there exists a $\delta \in \mathbb{N}$ such that the quantity

$$\sqrt[2h]{\frac{\hat{\alpha}_1^2(\hat{\alpha}_1 + \hat{\beta}_1)^{-2}}{1 - 2\hat{\alpha}_1\hat{\beta}_1 - \hat{\beta}_1^2}}$$

for $h \geq \delta$, works as a multiplicative bias correcting factor for the MLE $\hat{\phi}$.

Acknowledgements

Many thanks to an anonymous referee for his very useful comments on earlier draft of the manuscript.

References

[1] Bollerslev, T., Engle, R.F. and Nelson, D.B., (1994), ARCH Models, Ch. 49 [:in] *Handbook of Econometrics*, Elsevier Science, Amsterdam.

[2] Brockwell, P.J., Davis, R.A., (1987), *Time Series: Theory and Methods*, Springer-Verlag, New York.

[3] Hong, X. Billings, S. A, (1999), Time Series Multistep-Ahead Predictability Estimation and Ranking, *Journal of Forecasting* 18, 139–149.

[4] Hwang, S., Valls Pereira, P.L., (2006), Small Sample Properties of GARCH estimates and Persistence, *The European Journal of Finance* 12, 473—494.

[5] Pena, D. Sanchez, I., (2007), Measuring the Advantages of Multivariate versus Univariate Forecasts, *Journal of Time Series Analysis* 6, 886–909.

[6] Otranto, E., Triacca, U., (2007). Testing for Equal Predictability of Stationary ARMA Processes, *Journal of Applied Statistics* 34, 1091-1108.